The ArqusTermBot: A Prototype for the End-User (Short Paper)

Helena Ortiz-Garduño[*,†], Esther Castillo-Pérez[†], and Silvia Montero-Martínez[†]

University of Granada, C/Puentezuelas 55, 18002, Granada, Spain

Abstract

Terminology management is crucial for the internationalisation of academic and corporate institutions. The Arqus Termbase, a multilingual terminology database within the Arqus European University Alliance, exemplifies how academic institutions can efficiently manage terminology. To enhance resource accessibility, as a preliminary step, this study introduces a rule-based chatbot prototype. This tool enables users to query terminology and gain specific information about its usage and terminological concepts.

Keywords

Termbases, terminology management, chatbots, Arqus Termbase, user experience

1. Introduction

While corporate and governmental sectors boast numerous terminology databases, proactive management in the academic realm is scarce, as evidenced by the limited development of institutional terminology resources [1]. The Arqus European University Alliance stands out by initiating this process through the creation of the Arqus Termbase (ATB) prototype (https://arqusterm.ugr.es/) [2], a multilingual termbase. This resource is tailored to encompass terminology pertinent to the Arqus Alliance, its institutional partners, and European higher education.

Recent years have witnessed a notable surge in the adoption of chatbots (CBs) as conversational agents on websites across various sectors, including education, e-commerce, healthcare, and entertainment [3]. This trend has substantially altered the dynamics of how organisations engage with their customers and users. However, their utilisation in the domain of terminology management remains relatively constrained.

This study aims to create a preliminary, rule-based CB focused on enhancing access to information within ATB. The objective is to improve resource accessibility, enabling users to conduct terminological queries and seek guidance on tool usage.

3rd International Conference on "Multilingual digital terminology today. Design, representation formats and management systems" (MDTT) 2024, June 27-28, 2024, Granada, Spain

[*] Corresponding author.

[†] These authors contributed equally.

helenaortiz@ugr.es (H. Ortiz-Garduño); esthercaspe@ugr.es (E. Castillo-Pérez); smontero@ugr.es (S. Montero-Martínez)

0009-0001-4886-441X (H. Ortiz-Garduño); 0000-0002-1030-7693 (E. Castillo-Pérez); 0000-0002-2064- 401X (S. Montero-Martínez)

2. Termbases and user profiles

Terminology holds significant importance for communication and organisation in both public and private institutional settings. Effective terminology management has become essential, particularly in internationalised contexts, to bolster inter-institutional relations and enhance intra-institutional communication [4].

This encompasses tasks such as compiling, describing, documenting, and disseminating terms, along with resolving inconsistencies and developing terminological tools [5,6]. In complex institutional contexts, the creation of centralised terminology databases, called termbases (TBs) [7], is essential for efficiently managing, controlling, updating, and disseminating an institution's terminology. According to Granda & Warburton [8], the utilisation of these resources represents a significant improvement over dictionaries and glossaries, offering infinite possibilities for data consultation and structuring.

When designing and developing a TB, it is crucial to consider the users' needs [9,10,11]. In the Function Theory of Lexicography, Bergenholtz and Tarp [12] argue that lexicographic resources should be tailored to fulfil specific functions and address particular information needs. Expanding on this idea, Tarp [13] suggests that "users do not have specific needs unless they are related to specific types of situations." This implies that users' requirements differ depending on their usage context, categorised into communicative user situations and cognitive situations [14]. As López Rodríguez [15] rightly points out, this perspective on the significance of user needs can also be extended to terminographic resources.

In the case of the ATB, the "Arqus Questionnaire on Terminology Resources" [16] was conducted in 2020 to ascertain the requirements and expectations of the Arqus Alliance partners. The questionnaire enabled the compilation of all requirements for the ATB, encompassing languages, format, descriptive fields, data categories and user profiles. Specifically, among the Arqus Alliance member universities, the following end-user primary profiles were identified: administrative and academic personnel and students. Nevertheless, being an open resource publicly financed, it also contemplates external users, such as the general public.

3. Chatbots and user experience

CBs are sophisticated software systems designed to mimic human conversation, employing Natural Language Processing (NLP) to understand and respond to user inquiries [17]. Acting as intelligent entities, these digital assistants streamline interactions between the user and the machine by providing timely and relevant responses, thereby enhancing user experience across various platforms.

In a world that prioritises efficiency and speed, CBs have emerged as a tool capable of delivering quicker and more effective responses than traditional manual systems [18]. Moreover, CBs enable uninterrupted, continuous service and offer the advantage of gathering information about users making queries. This allows organisations to consistently enhance service quality and provide pertinent answers thanks to (1) rule-based CBs, employing a predefined set of rules to respond to users; and (2) neural network-

based CBs, capable of generating responses either from a vast dataset (retrieval-based models) or by formulating responses from scratch (generative models) [19].

In the realm of higher education, CBs have the potential to enhance interpersonal communication, learning experiences, and the delivery of diverse information and knowledge, given their interactive and intuitive nature [20]. A growing number of academic institutions are integrating CBs to improve user experiences. In the case of the University of Granada, the CB Alhe has been implemented to assist students and prospective students by facilitating inquiries related to educational offerings, access, admission, pre-registration, academic procedures, international mobility, and scholarships, among others.

However, their utilisation in the domain of terminology management, particularly in connection with TB usage, remains constrained. For the purpose of this preliminary work, the rule-based architecture for the ArqusTermBot focuses on enhancing access to information within ATB. To cater to end-user profiles who lack specific knowledge of terminology and the utilisation of TBs, the final objective is to enable users to conduct terminological queries and seek guidance on tool usage.

4. Materials and methods

This study is founded on the fundamental premise that the incorporation of CBs enhances the end-user experience with institutional TBs. The objective of this preliminary research is to enhance ATB by creating a dedicated CB prototype to streamline information access. The subsequent sections elaborate on the key attributes of the multilingual ATB, the utilisation of the Landbot.io platform for CB design and development, and the various stages involved in the creation of ArqusTermBot.

4.1. Arqus Termbase

Arqus Termbase is the centralised database of the Arqus European University Alliance (https://arqusterm.ugr.es/), a multilingual network of European academic institutions. This resource, tailored to the needs of the Alliance [21], endeavours to enhance communication coherence and fluidity, through centralised terminology management, elimination of inconsistencies, and information dissemination regarding term usage [22].

ATB (Figure 1) presently incorporates seven official languages from partner universities, and access to terminology is provided from term to concept (semasiological access) and from concept to term (onomasiological access). It features a simple and advanced approximate search system [23], which accommodates diverse consultation paths based on the profile and requirements of the end-user.

In designing ATB, primary consideration was given to the diverse users within the university communities of Arqus Alliance members. They include administrative staff, teaching and research personnel, and students. Nevertheless, being an open resource, it also targets external users, including translators, interpreters, interested institutions, and the general public. Consequently, user-friendly help sections must be developed to cater to user profiles lacking specific knowledge of terminology and the utilisation of normative resources.

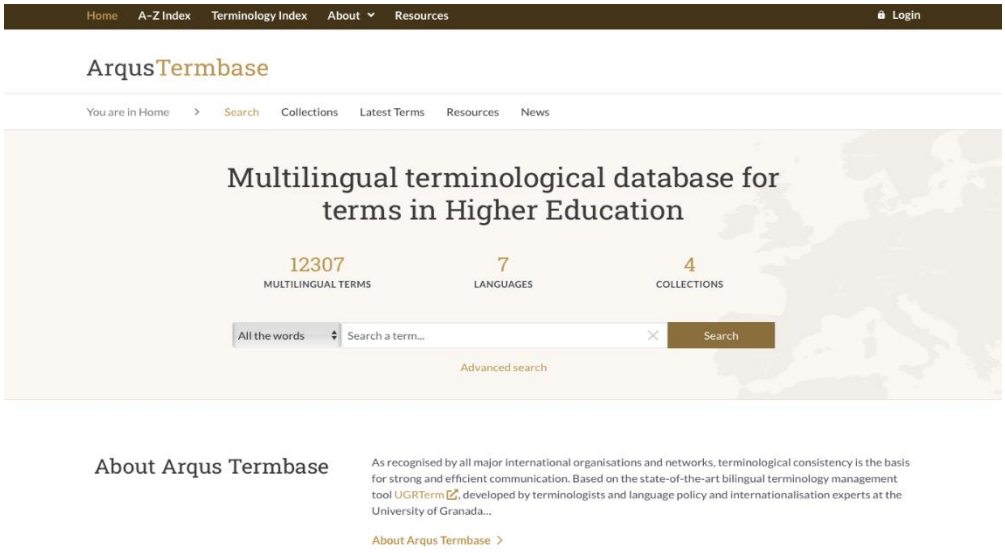

Figure 1: ATB homepage

4.2. The Landbot.io platform

The CB prototype was constructed using Landbot.io (https://landbot.io/) [24], a tool that enables the intuitive development and customisation of conversational experiences without the necessity for coding. The CB interface is organised into five sections.

The "Build" section constitutes the central component of the CB design, offering a platform to construct and configure the logic of conversations. This section delineates dialogue flows, actions based on user interactions, and the format of questions and answers. A distinction is made between open-ended questions, allowing the use of natural language (e.g. the user's profile and the institution to which they belong), and multiple-choice questions, where users simply click on the desired topic (e.g. information on the structure of ATB, the languages it covers, the type of terminology it contains, and how it can be accessed).

The "Design" section concentrates on the visual aesthetics of the CB, offering tools to customise its interface and align it with the visual identity of the intended product. Customisation options encompass font type and size, background design, logo selection, and personalisation of the virtual assistant's messages and buttons.

Within the "Settings" section, various options enable the comprehensive configuration of the virtual assistant, ranging from the CB's name to various system customisation parameters. This includes the ability to modify default messages in text insertion segments, help buttons, question fields, or error messages. Additionally, features such as activating the type emulator and setting the average human reading speed are available to enhance the user experience, making it as close as possible to interacting with a real person.

The "Share" section offers various options for sharing and previewing the designed CB. Users can choose how the virtual assistant is displayed, such as a pop-up conversation on the right-hand side of the screen or as a full web page. Moreover, the option to share the CB via a URL enables testing the designed flow without making it publicly accessible to all users.

Lastly, the "Analyze" section focuses on monitoring and evaluating the tool's performance. It offers metrics on user interaction and a comprehensive analysis of the flow, including percentages indicating the path taken by users. Furthermore, user responses are categorised based on the predefined variants, simplifying data collection. This allows CB creators to iteratively optimise the tool, enhancing its effectiveness and aligning it with observed needs from the analysis.

4.3. Chatbot design phases

The design of the CB is crucial in establishing the groundwork to meet users' expectations, aiming to address their needs quickly and effortlessly compared to alternative solutions [25]. In this instance, a set of guidelines was formulated based on recommendations from various authors to ensure the functionality and effectiveness of the tool [26, 27].

The initial phase involved comprehensive planning, wherein the objectives of the virtual assistant were delineated, users' needs identified, and the purpose of interactions established. Additionally, the personality of the CB was defined based on the anticipated user interactions. Subsequently, the conversational flow was meticulously designed. A set of questions and their corresponding answers were developed to address the most frequently asked questions anticipated from users. These questions were organised into categories, and a flow diagram was created using the Landbot.io platform. The diagram visually represented various bifurcations based on user responses, outlining the rules that define the CB's architecture. Additionally, specific variables were established to collect user-provided information, while maintaining data privacy. During this phase, a significant challenge involved the difficulty in anticipating and covering the full range of user queries, which can vary widely in scope and specificity. The inherent constraints of a rule-based architecture limit the CB's response flexibility, confining interactions to a pre-established set of conditions and pathways.

The iterative refinement phase involved pilot tests to validate functionality and identify errors. Issues related to level structure and conversation flow were addressed, with options allowing users to return to higher levels without re-entering data. Decision points posing a risk of user dropout were simplified, and closed answers were prioritised over open-ended questions to avoid overwhelming users with excessive natural language. Finally, the visual identity of the CB was crafted to enhance the user experience, drawing inspiration from the existing ATB. The primary challenge of this phase was enhancing the CB's structural fluidity to prevent user frustration and potential abandonment. However, a limitation faced is the inability to precisely ascertain the dropout risk, as this CB prototype has not yet been launched publicly, thereby constraining the ability to analyse and address user behaviour patterns effectively before its release.

5. The ArqusTermBot prototype

To fulfil the requirements associated with information retrieval in ATB, this preliminary study has chosen to develop a rule-based CB prototype. This architectural approach facilitates logical and efficient responses to the queries posed by ATB users, aligning well with the nature of their questions. Consequently, ArqusTermBot is crafted as a virtual

assistant with a friendly and empathetic personality, encouraging users to stay engaged until their inquiries are resolved.

The conversational flow is structured into two primary rule groups. The first addresses queries regarding the utilisation of the terminology tool, and resolves doubts about ATB's structure, terminology access, term typology, languages, entry fields and tool managers. The second elucidates fundamental concepts in the field of terminology for different user profiles, who can seek answers related to the definition of key concepts such as "terminology," "term" and "terminology entry," as well as understand the relationships between them (Figure 2). The goal is to provide users of ATB with the opportunity to acquire a basic understanding of terminology before actively using the resource, ensuring a practical and accurate application of the terminology tool.

Figure 2: ArqusTermBot internal structure on terminology information

Regarding the visual identity of the CB, as depicted in Figure 3, a design aligned with the ATB has been selected. Consequently, the name and visual image incorporate the Alliance logo and the corporate colour palette.

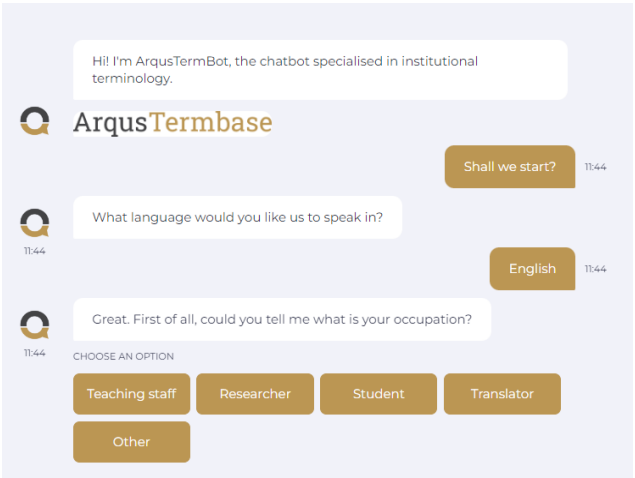

Figure 3: ArqusTermBot design interface

Concerning data collection, the CB is equipped to record data from users interacting with the tool. Variables set during the development phase enable efficient data collection. For instance, the CB can capture information about the user type (teaching and research staff, administrative staff, students, translators, or other) making queries, and the institution to which they belong (Figure 2). However, to uphold confidentiality, the interface refrains from collecting personal data. Instead, it indicates user registration through an internal identifier. In instances of specific inquiries or improvement suggestions, the CB redirects users to the contact section of ATB.

As a result, the incorporation of the ArqusTermBot within ATB will significantly enhance user engagement and information accessibility, especially for end-users with no previous experience using multilingual terminology resources. By streamlining the querying process, the CB reduces the time and effort required for users to handle the TB in order to perform relevant and appropriate searches. Moreover, the CB's ability to provide immediate, accurate responses fosters an interactive and engaging way to navigate terminology content. Therefore, ArqusTermBot represents a significant step towards a more accessible and efficient ATB, not only for the academic community but for other stakeholders.

Notwithstanding, further developments and improvements are considered for this prototype: (1) expanding the rules governing the CB to cover a broader spectrum of questions that users may pose; (2) including the languages of Arqus' member institutions to have a multilingual CB; (3) investigating the development of a neural network based version, which could address several limitations inherent to rule-based CBs, and (4) incorporating metrics analysis to conduct a more comprehensive evaluation of user interaction, which will involve tracking flow analytics, response accuracy, and user satisfaction levels to identify patterns and areas for improvement.

6. Conclusions

The design of a rule-based CB prototype for ATB serves a dual purpose. Firstly, it acts as an intuitive tool, aiding users in becoming more familiar with this institutional terminology resource. Secondly, it facilitates access to terminological knowledge by addressing the terminology-related inquiries of users interacting with ATB. The benefits of this research extend to the Arqus European University Alliance partner community and anyone interested in applying these technological capabilities to the context of terminology management and TBs.

However, to address a broader spectrum of end-users, this prototype requires further development to overcome issues such as ArqusTermBot's inability to handle unexpected user inputs and confinement to predefined rules. This will enhance the CB's ability to understand and respond dynamically to complex queries, thereby improving adaptability and response accuracy.

Acknowledgements

This work was supported by the Spanish Ministry of Universities [*Predoctoral Grants for the Training of University Lectures (FPU), FPU21/01204 and FPU20/03089*], the European

Commission [*Arqus European University Alliance, 612247-EPP-1-2019-1-ES-EPPKA2-EUR-UNIV*], the European Commission Erasmus + European Universities [*Arqus II, ERASMUS-EDU-2022-EURUNIV-1*], and the Spanish Ministry of Science and Innovation [*Integración transversal de la cultura en una base de conocimiento terminológico medioambiental, PID2020-118369GBI00*].

References

[1] E. Castillo-Pérez, H. Ortiz-Garduño, S. Montero-Martínez, Las bases de datos terminológicas y la gestión institucional: Utilidad y limitaciones, in: Arqus Symposium on Translation and Interpreting Studies Translation and Interpreting: Bridging the Gap between the Classroom and the Profession (1ª Ed.), Universidad de Padua (Italia), 2022, pp. 5. URL: http://www.maldura.unipd.it/Arqus-tt-2021/#detail.

[2] Arqus European University Alliance, Arqus Termbase, 2023. URL: https://arqusterm.ugr.es

[3] G. Caldarini, S. Jaf, K. McGarry, A literature survey of recent advances in chatbots, Information 13(1) (2022) 41-63. doi:10.3390/info13010041.

[4] E. Castillo-Pérez, S. Montero-Martínez, Internationalization and terminology management in higher education, in: Book of abstracts of the Linguistic, Educational and Intercultural Research Conference, LEIC 2021, Vilniaus universitetas, Vilna, 2021, pp. 72–73. URL: https://www.conference.uki.vu.lt/images/2021/The_book_of_abstracts_LEIC_conference_2021.pdf.

[5] TerminOrgs, Terminology Starter Guide, K. Warburton (Ed.), 2016. URL: https://www.terminorgs.net/downloads/TerminOrgs_StarterGuide_V2.pdf.

[6] R. Granda, K. Warburton, Terminology Management as Data Management, in: Proceedings of the 2001 Conference of the Centre for Advanced Studies on Collaborative Research, CASCON '01, IBM Press, 2001, pp. 4.

[7] ISO 12616-1, Terminology working support of multilingual communication – Part 1: Fundamentals of translation-oriented terminography (Patent ISO 1261-1:2021), International Organization for Standardization, 2021.

[8] R. Granda, K. Warburton, Terminology Management as Data Management, in: Proceedings of the 2001 Conference of the Centre for Advanced Studies on Collaborative Research, CASCON '01, IBM Press, 2001, pp. 4.

[9] H. Bergenholtz, S. Tarp, The concept of dictionary usage, Nordic Journal of English Studies (2004) 23- 36.

[10] S. Tarp, Lexicography in the information age, Lexikos 17 (2007) 170-179.

[11] C. I. López Rodríguez, M. Buendía Castro, A. García Aragón, User needs to the test: evaluating a terminological knowledge base on the environment by trainee translators, The Journal of Specialised Translation 18 (2012) 57-76.

[12] H. Bergenholtz, S. Tarp, The concept of dictionary usage, Nordic Journal of English Studies (2004) 23- 36.

[13] S. Tarp, Lexicography in the information age, Lexikos 17 (2007) 170-179.

[14] S. Tarp, Lexicography in the information age, Lexikos 17 (2007) 170-179.

[15] C. I. López Rodríguez, M. Buendía Castro, A. García, User needs to the test: evaluating a terminological knowledge base on the environment by trainee translators, The Journal of Specialised Translation 18 (2012) 57-76.

[16] S. Montero-Martínez, P. León-Araúz, M. García de Quesada, Arqus Questionnaire on Terminology Resources, 2020 [Available upon reasonable request].
[17] Comisión Mundial de Ética del Conocimiento Científico y la Tecnología (COMEST) de la UNESCO, Preliminary study on the ethics of artificial Intelligence, UNESCO, París, Francia, 2019. URL: https://unesdoc.unesco.org/ark:/48223/pf0000369455.
[18] A. Rane, C. Ranade, H. Bandekar, R. Jadhav, V. Chitre, AI driven Chatbot and its Evolution, in: 5th International Conference on Advances in Science and Technology (ICAST), Bombay, India, 2022, pp. 170-173. doi: 10.1109/ICAST55766.2022.10039515.
[19] R. Agarwal, R., M. Wadhwa, Review of State-of-the-Art Design Techniques for Chatbots, SN Computer Science 1(5) (2020). doi: 10.1007/s42979-020-00255-3.
[20] G. J.Hwang, C. Y. Chang, C. Y., A review of opportunities and challenges of chatbots in education, Interactive Learning Environments 31(7) (2021) 4099–4112. doi: 10.1080/10494820.2021.1952615.
[21] S. Montero-Martínez, P. León-Araúz, M. García de Quesada, Arqus Terminology Questionnaire: Quantification and Analysis, 2020. https://arqusterm.ugr.es/resource/arqus-terminology-questionnaire-quantification-and-analysis-2020/.
[22] S. Montero-Martínez, E. Castillo-Pérez, Arqus Termbase: Terminology for the Arqus Community, in: 3rd Wrocław Terminological Meeting, TERMOS 2023 "Terms and Their Users", 2023.
[23] S. Montero-Martínez, E. Castillo-Pérez, Arqus Termbase: Multilingual Terminology for the Arqus European University Alliance, in: Languages in European University Alliances: challenges and issues, Conseil Européen pour les Langues/ European Language Council, 2022.
[24] Hello UMI S.L., Landbot.io, 2023. URL: https://landbot.io/.
[25] S. Bisser, Microsoft Conversational AI Platform for Developers: End-to-End Chatbot Development from Planning to Deployment, Springer, New York, NY, 2022.
[26] A. Campos-Magencio, Chatbots in Education, in: Chatbots, inteligencia artificial y mensajería instantánea, Hello UMI (Madrid), 2017. URL: https://www.youtube.com/watch?v=cKSeipwPW-U.
[27] S. Bisser, Microsoft Conversational AI Platform for Developers: End-to-End Chatbot Development from Planning to Deployment, Springer, New York, NY, 2022.